Accelerated Strategic Computing Initiative

--

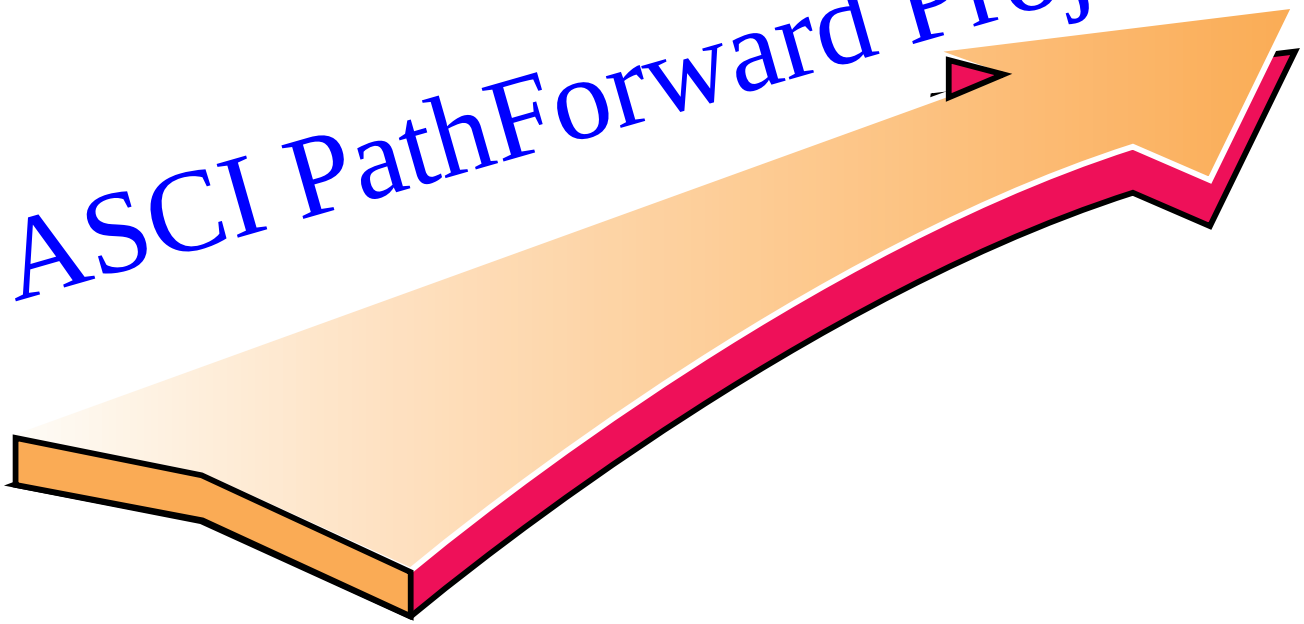
applications driven high-end computing

Gil Weigand

Deputy Assistant Secretary for Strategic Computing and Simulation

U.S. Department of Energy

ASCI PathForward Project



President Clinton's Vision:

"...we can meet the challenge of maintaining our nuclear deterrent under a [comprehensive test ban] through a stockpile stewardship program without nuclear testing."

Background on the Accelerated Strategic Computing Initiative (ASCI)

On August 11, 1995 President Clinton announced the United States' intention to pursue a "zero-yield" Comprehensive Test Ban Treaty:

> "...we can meet the challenge of maintaining our nuclear deterrent under a [comprehensive test ban] through a science-based stockpile stewardship program without nuclear testing."
>
> - President Clinton

As a response to this presidential mandate, ASCI was created with a vision to:

Shift promptly from nuclear test-based methods to computation-based methods. ASCI will create the leading-edge computational modeling and simulation capabilities that are essential for maintaining the safety, reliability, and performance of the U.S. nuclear stockpile and reducing nuclear danger.

ASCI Vision

Create leading-edge computational modeling and simulation capabilities critically needed to

promptly shift from nuclear test-based methods to computational-based methods,

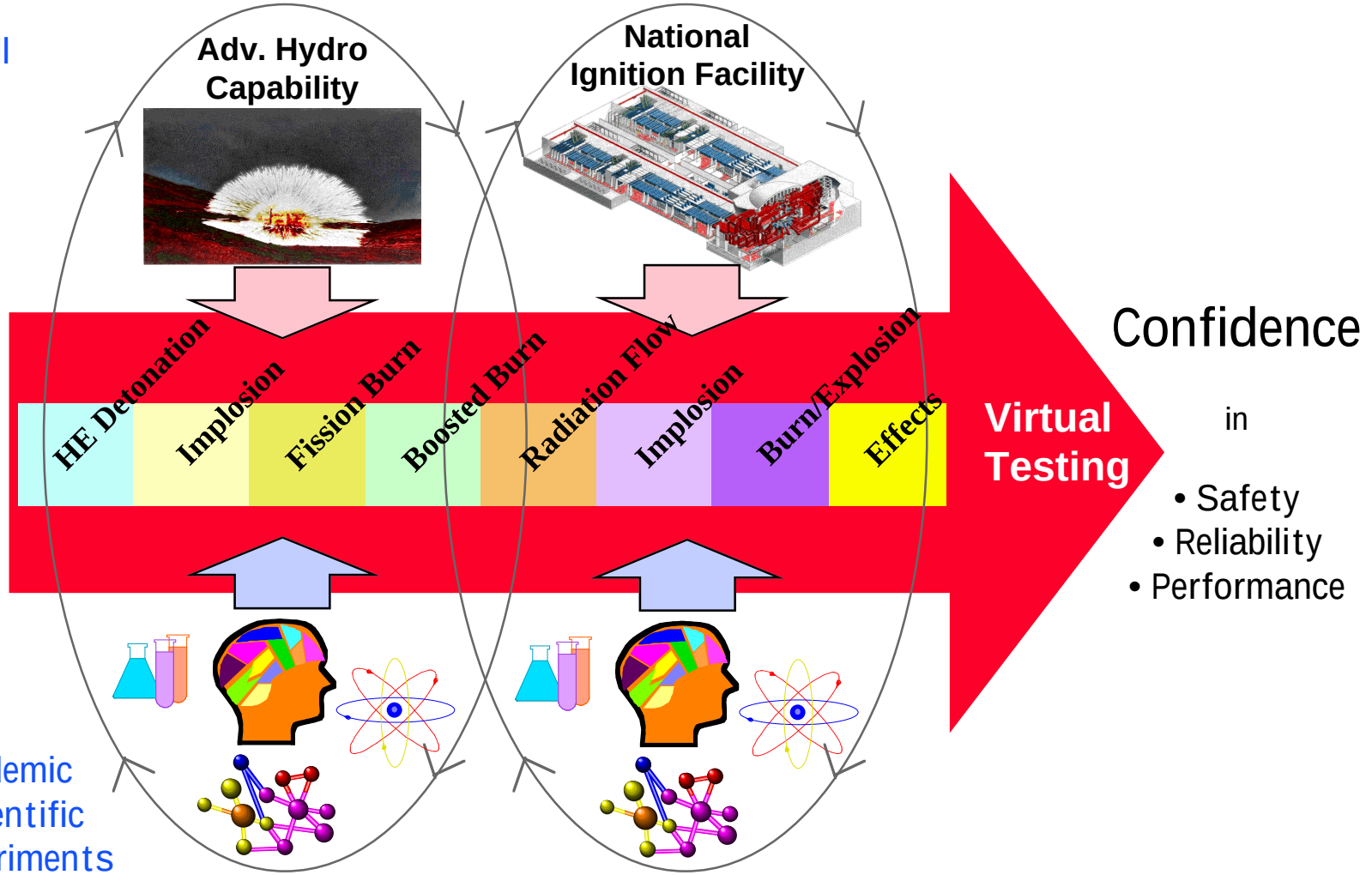
to integrate stockpile stewardship elements and provide an integrated nuclear explosion testbed.

Simulation Tools Provides Integration of Great Science, Experimental Facilities and Archive Data for Confidence in the Stockpile

① Full scale or full energy-density studies & experiments

② Computation & Modeling

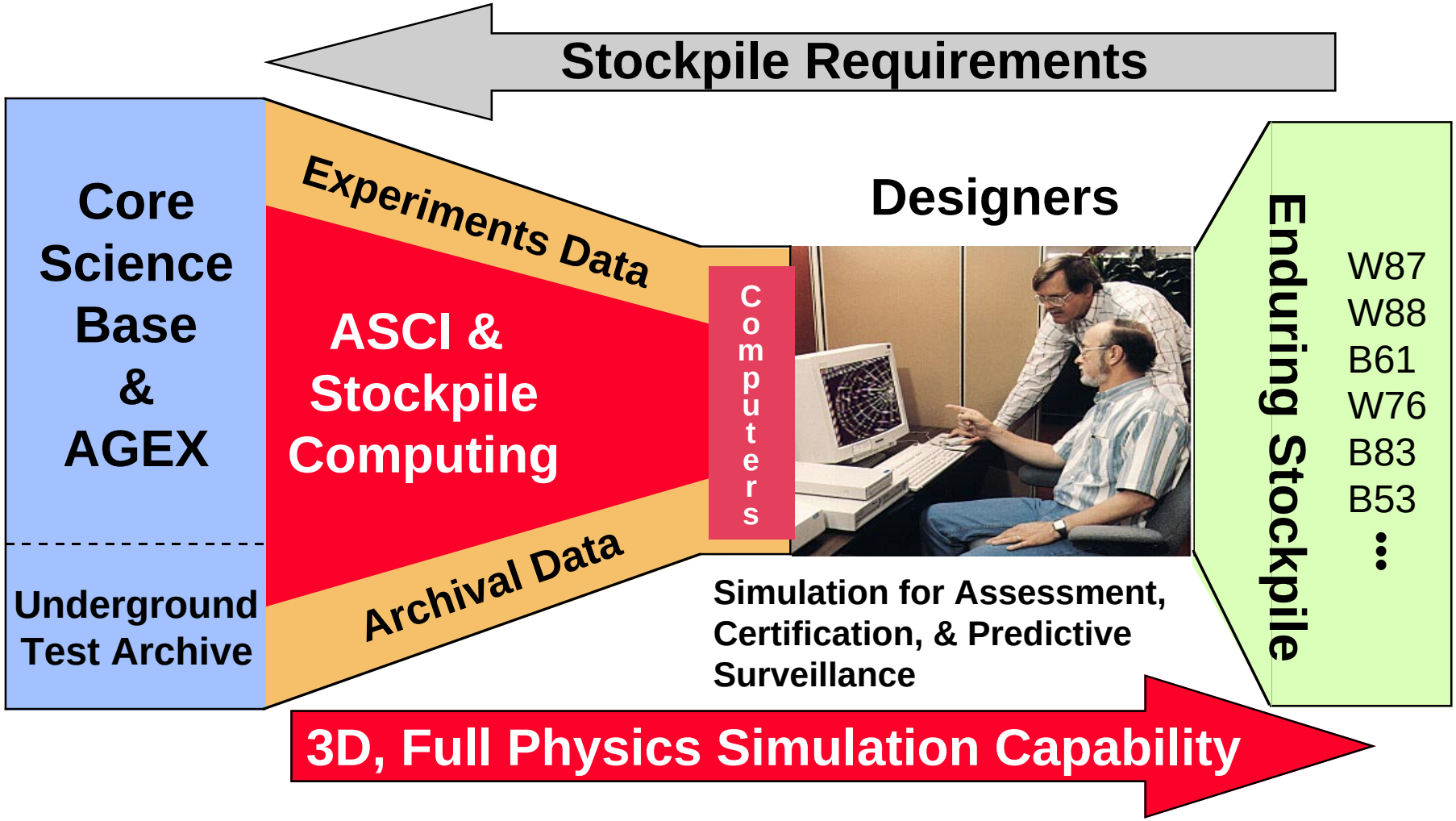
③ World-class & forefront, academic & lab scale scientific studies & experiments



Confidence in

- Safety
- Reliability
- Performance

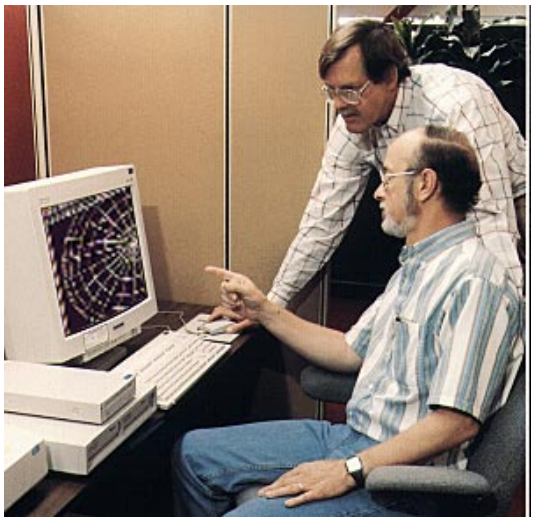
ASCI produces and supports simulation capabilities for designers who make critical decisions about the enduring stockpile



We are focused on Nuclear Weapons Simulation

Simulation

"digital proxy for physical"

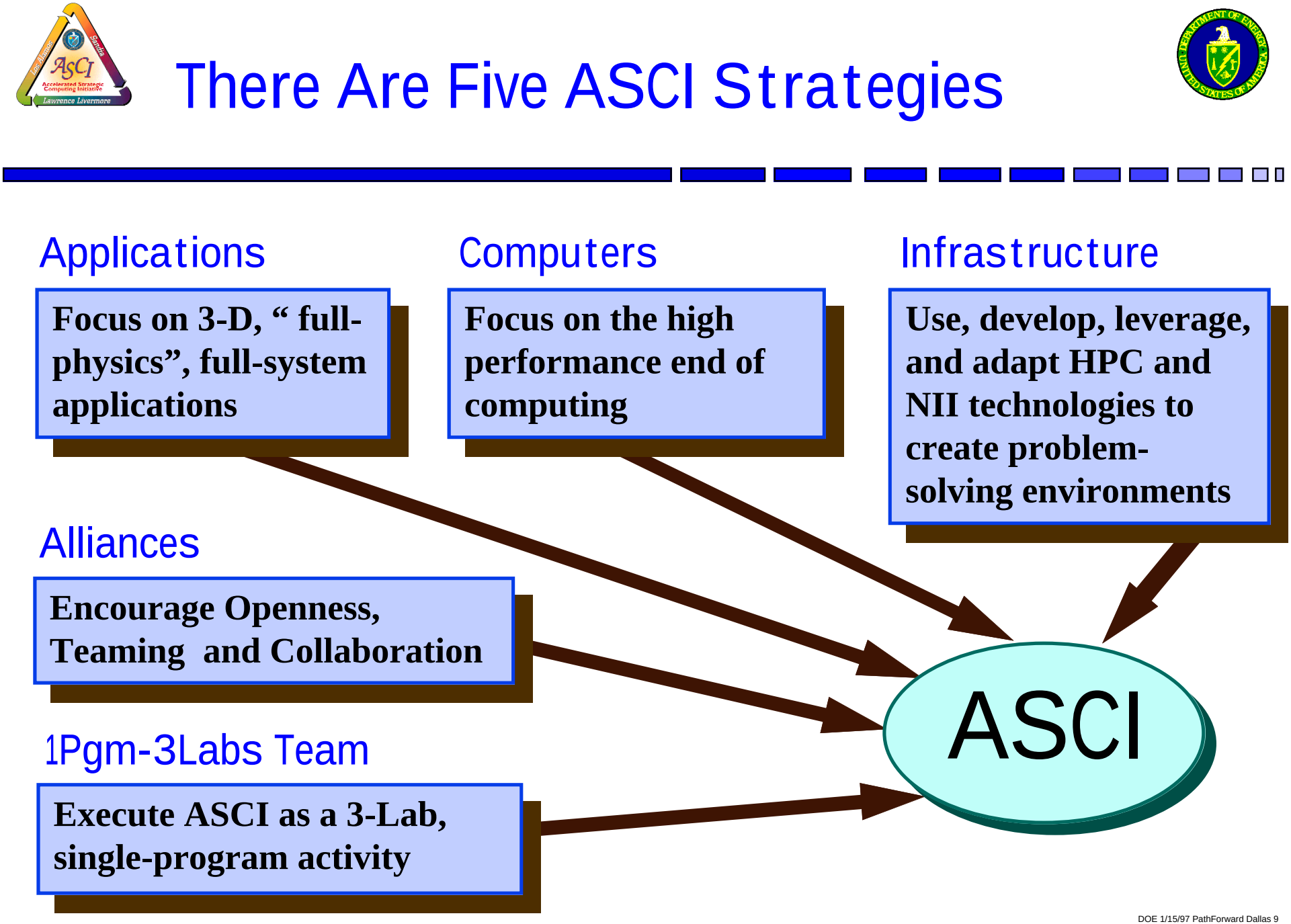


National Labs — applications, weapon science, simulation & AGEX, adv. computation

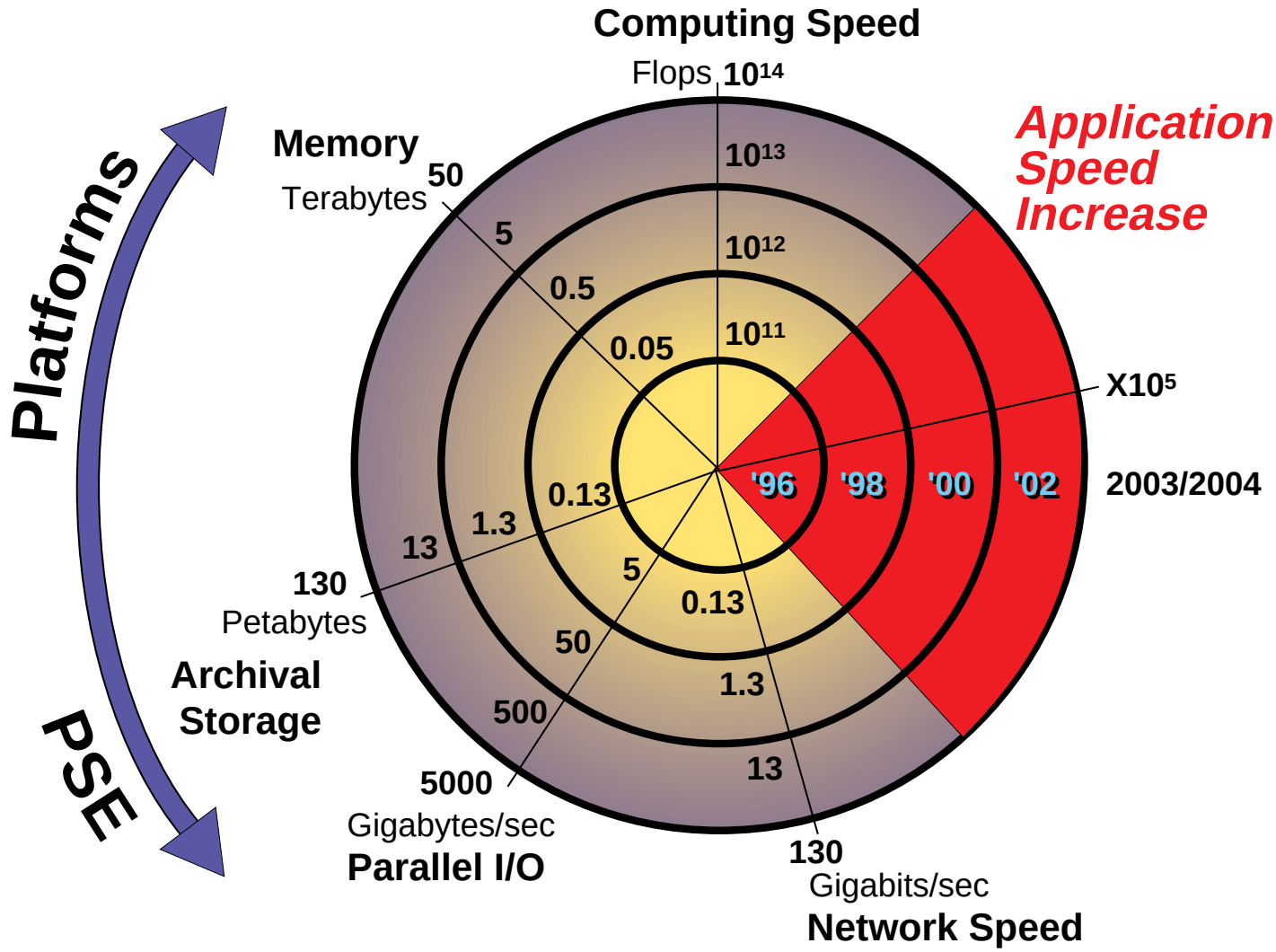
Computer Co's — computers, software & OS, prog. support

University & other Lab — science, collaboration, applications & prog. tools

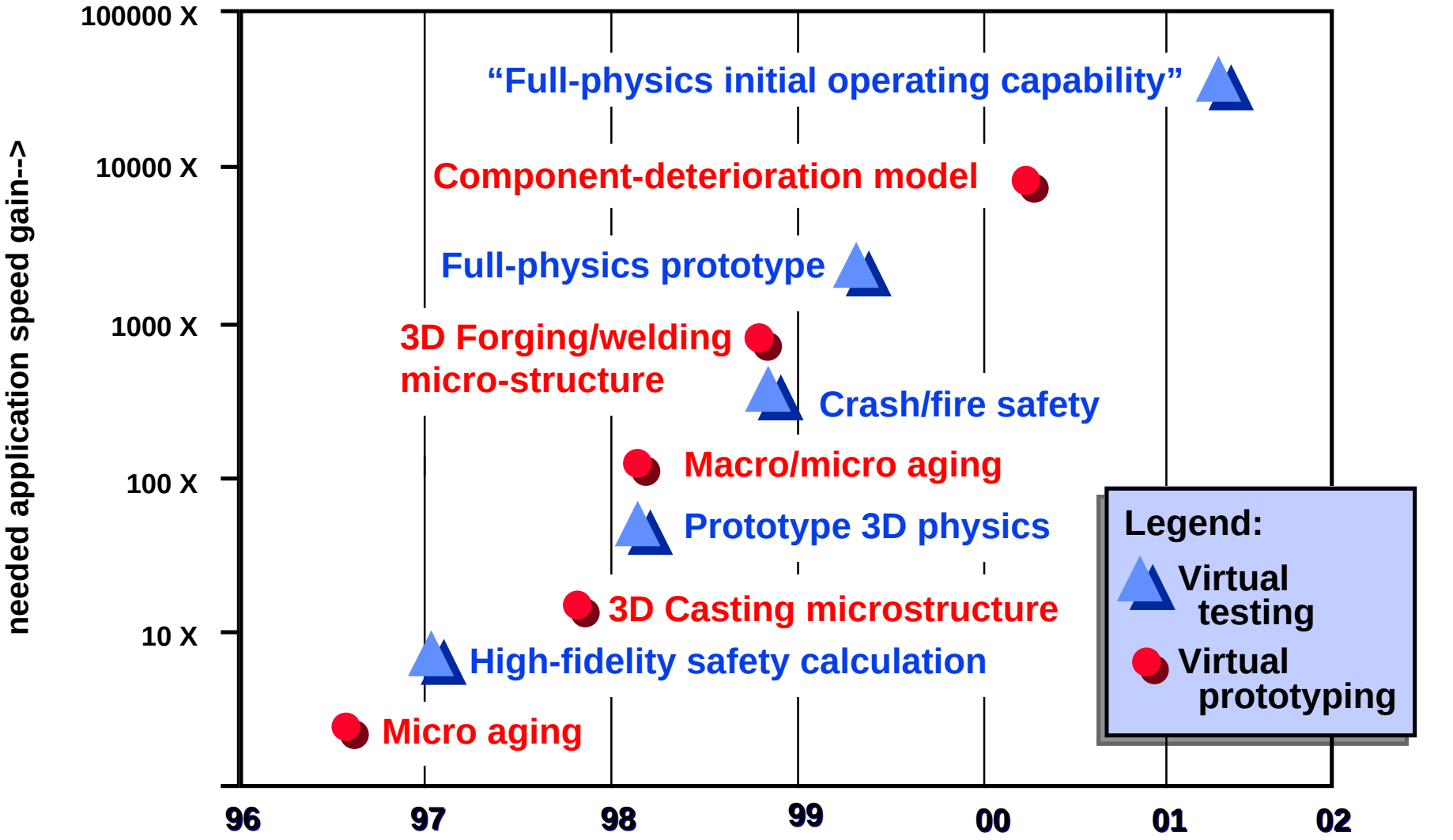
There Are Five ASCI Strategies



ASCI Success Depends upon Balanced Growth across all Strategies

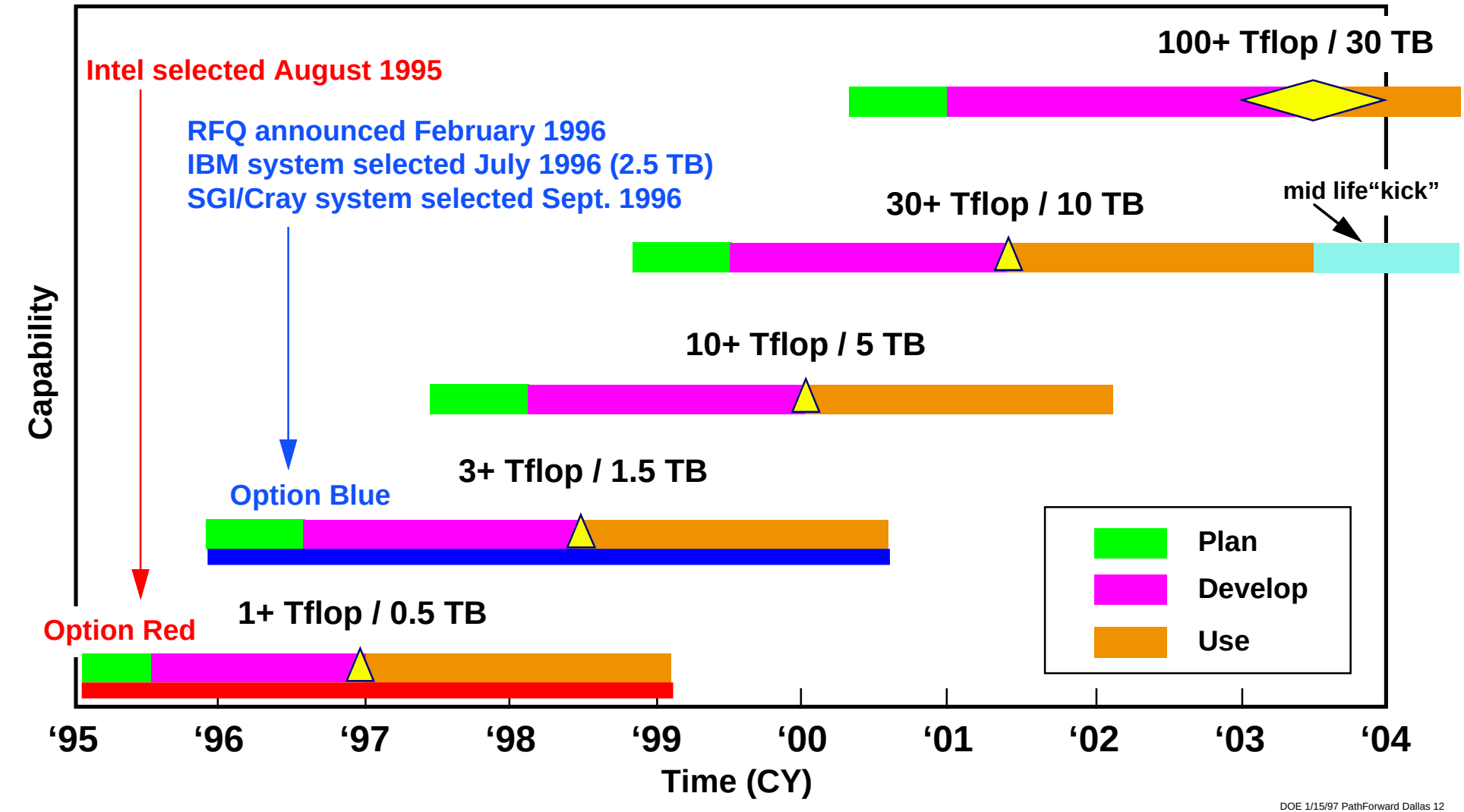


3D, "full-physics", high fidelity simulation, have key milestones tied to stockpile requirements and requires unprecedented application and computer speed increases

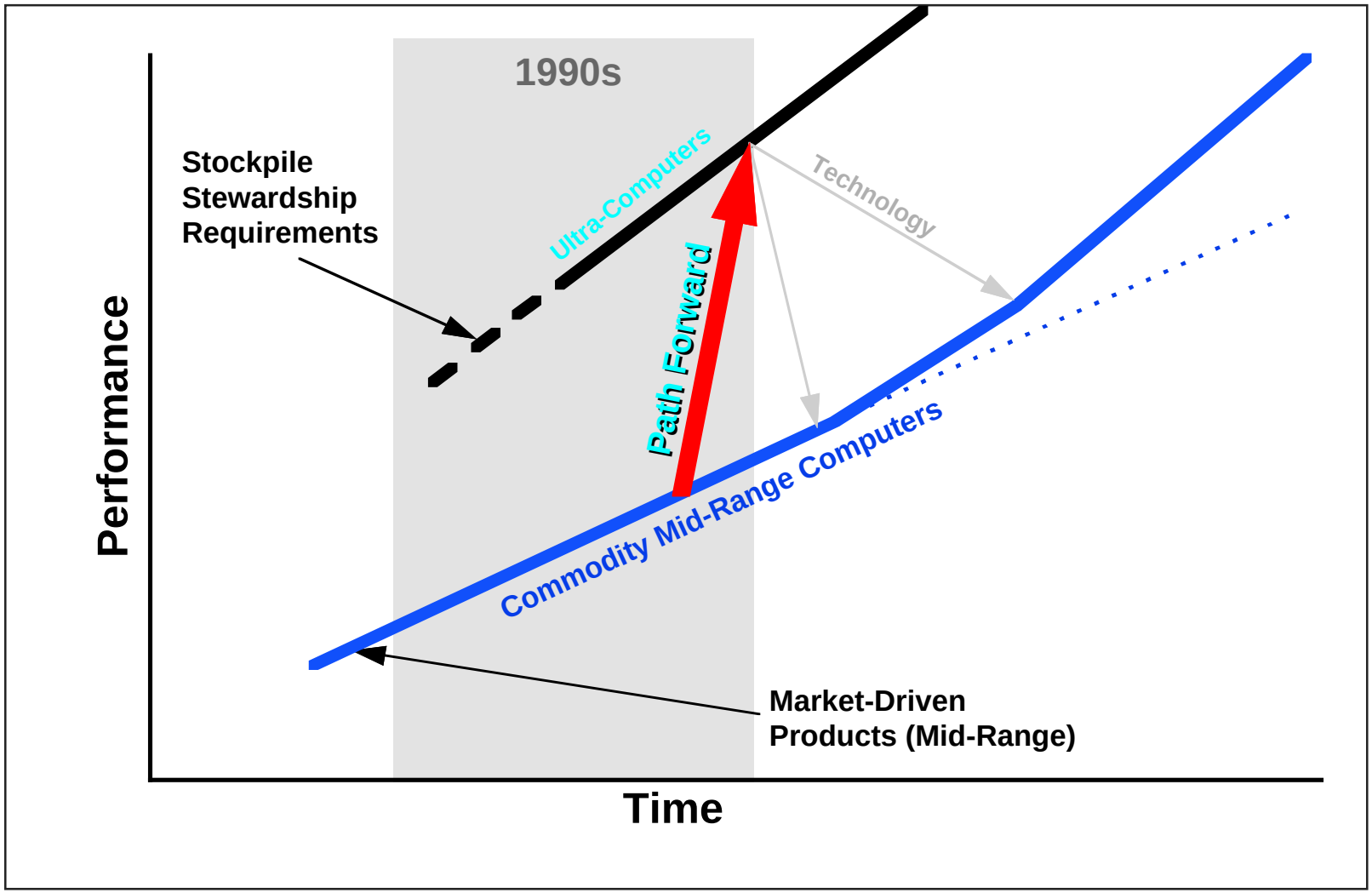


ASCI Computing Systems Roadmap

-- working with industry to reach unprecedented computer performance --



HEC Industry is not on the Path to Meeting Stockpile Stewardship Requirements



ASCI is Focused on the High End of Computing

- Accelerate High-Performance Computing through Multiple-Partnerships
- Develope Software to take full advantage of hardware capabilities
- Maintain affordability by acclerating existing industry technology trends
- Provide balanced 100 TFlops computing platforms by 2003/4

National security hardware requirements

Level	Effective Latency (CPU cycles)	Bandwidth (Random read/write)	Size	
On-chip cache**, L1	2-3 ●	16-32 B/cycle ●	10^{-4} B/flop * ●	
Off-chip cache**, L2 (SRAM)	5-6 ●	16 B/cycle ●	10^{-2} B/flop * ●	
Local main memory (DRAM)	30-80 (15-30)	2-8 B/flop pk (2-8 B/flop sustained)	1 B/flop ●	Compute engine
"nearby nodes"	300-500 (30-50)	1-8 B/flop (8 B/flop)	1 B/flop ●	Interconnect
"far away nodes"	1000 (100-200)	1 B/flop (1 B/flop)	1 B/flop ●	
I/O (memory disk)	10 ms ●	0.01-0.1 B/flop ●	10-100 B/flop ●	
Archive (disk-tape)	Seconds ●	10^{-4} B/flop (0.001-0.01 B/flop)	10^{2} B/flop 10^{4} B/flop	
User access	1/10 s (1/60 s)	OC3/desktop (OC12-48 /desktop)	100 users ●	
Multiple sites	1/10 s ●	●	●	

* Equivalent integer and floating-point data calculation rates are required.

** Cacheless systems with equivalent performance are fully acceptable.

Primary investment priority

Secondary investment priority

1996-1998 Situation

(1998-2000 Requirements)

Industry Trend

- Industry gets better at meeting requirements
- Industry gets worse at meeting requirements
- ● Industry continues to meet requirements

Current Hardware Trends (Commodity)

- industry, alone, will not meet national security requirements -

Level*	Effective Latency (CPU cycles)	Bandwidth (Random read/write)	Size
On-chip cache**, L1	•	•	↑
Off-chip cache**, L2 (SRAM)	•	•	↑
Local main memory (DRAM)	↓	↓	↑
Compute engine			
"nearby nodes"	↓	↓	•
Interconnect			
"far away nodes"	↓	↓	•
I/O (memory disk)	•	•	•
Archive (disk-tape)	•	↓	↓
User access	•	↓	•
Multiple sites	•	•	•

Primary investment priority

Secondary investment priority

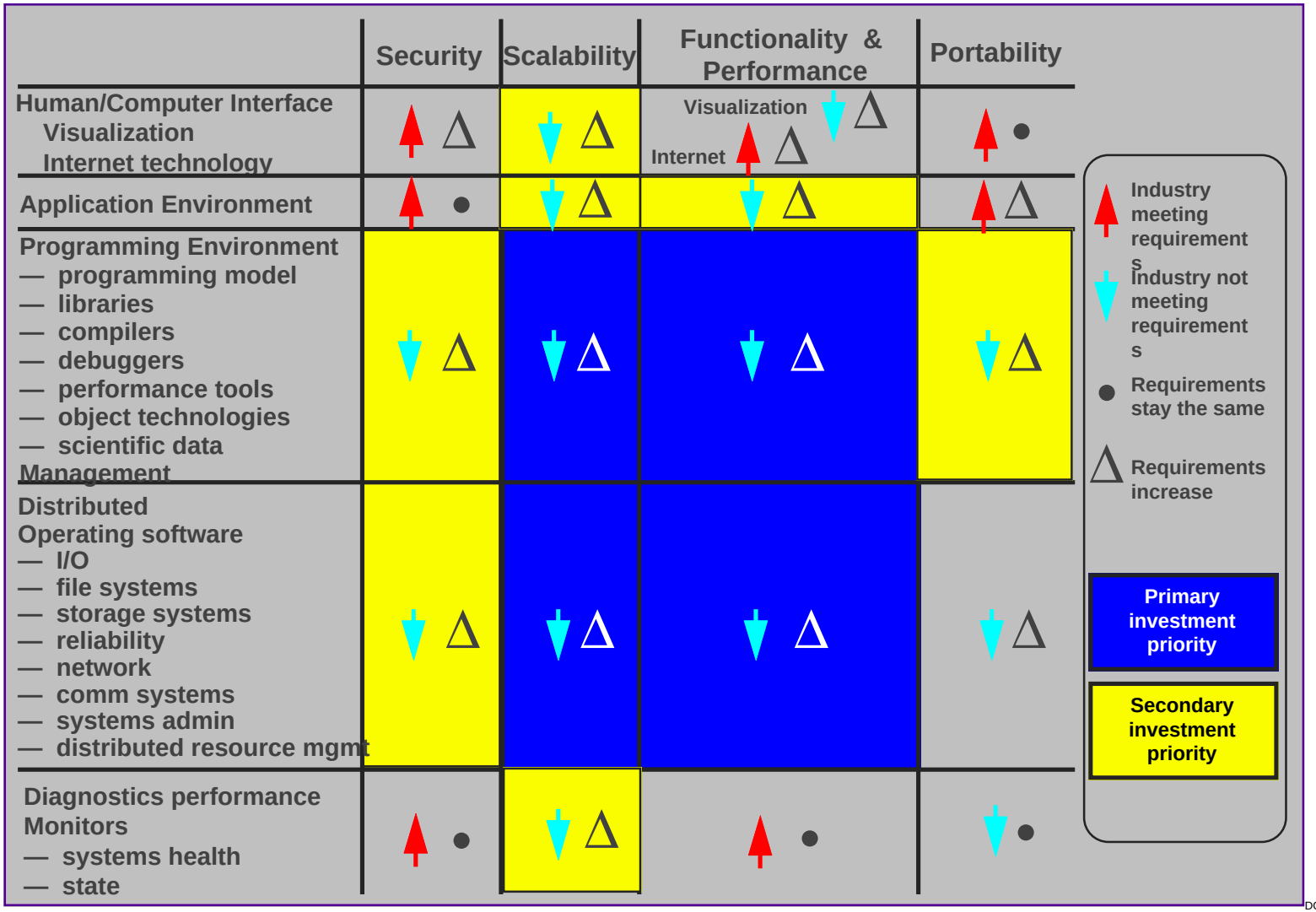
Industry Trend 1996-2000

- ↑ Industry gets better at meeting requirements
- ↓ Industry gets worse at meeting requirements
- • Industry continues to meet requirements

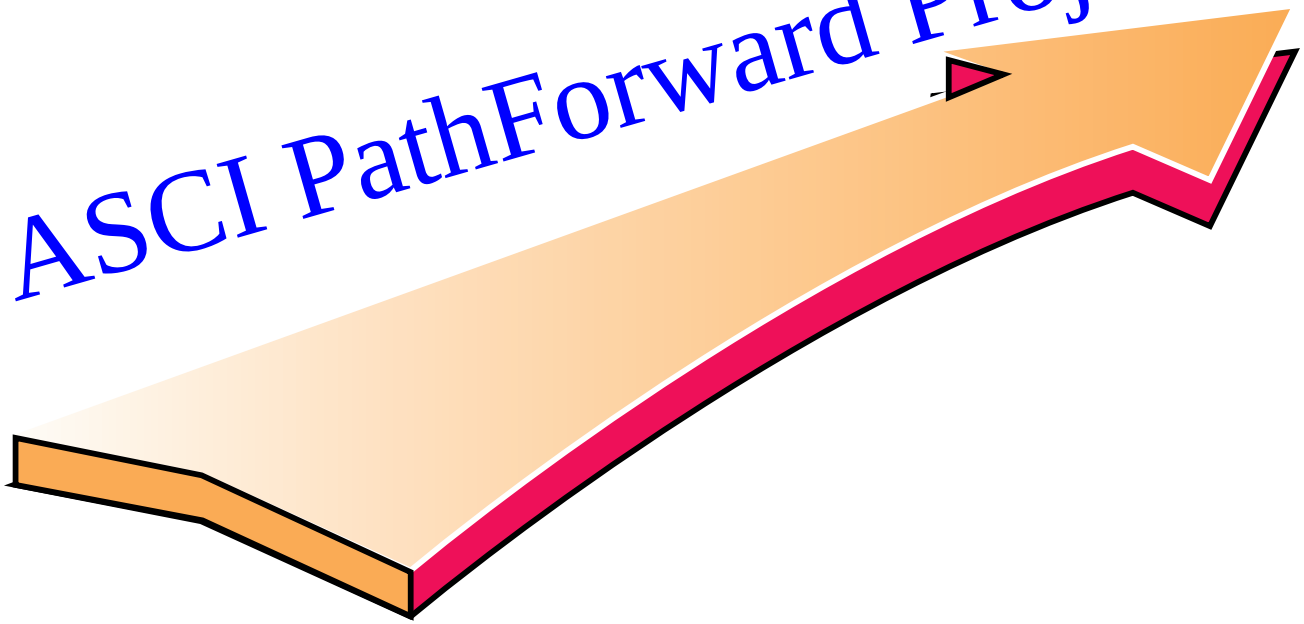
* Equivalent integer and floating-point data calculation rates are required.

** Cacheless systems with equivalent performance are fully acceptable.

National security software requirements



ASCI PathForward Project



PathForward Project

"Establish multi-year development and engineering efforts to create the critical integrating and scaling technologies required for balanced computing environments at the scale of 10 to 30 TeraFLOPS in the late 1999 to 2001 timeframe."

PathForward Project

Lawrence Livermore National Laboratory

Los Alamos National Laboratory

Sandia National Laboratory

U.S. Department of Energy, Defense Programs

PathForward Agenda

10:00-10:05	Welcome and Opening Comments	P. Smith
10:05-11:00	Science Based Stockpile Stewardship/ Accelerated Strategic Computing Initiative	G. Weigand
11:00-11:30	PathForward Project Overview	R. Deri
11:30-12:00	PathForward Interaction Process	D. Hammer
12:00-1:00	Lunch	
1:00-2:40	Panel Session and Participants Discussion on: Panel Development and Engineering Opportunities	
	- Interconnect Technology (20 min.)	R. Deri
	- Operating Systems (20 min.)	A. Hale
	- Programming Environment (20 min.)	K.-H. Winkler
	- Ultra-Scale Storage Technology (20 min.)	D. Watson
2:40-3:00	Open Question & Answers	Project Team
3:00	Adjourn	P. Smith

PathForward Project Team

http://www.llnl.gov/asci-pathforward

Bob Deri, LLNL

Art Hale, SNL

Derrol Hammer, LLNL

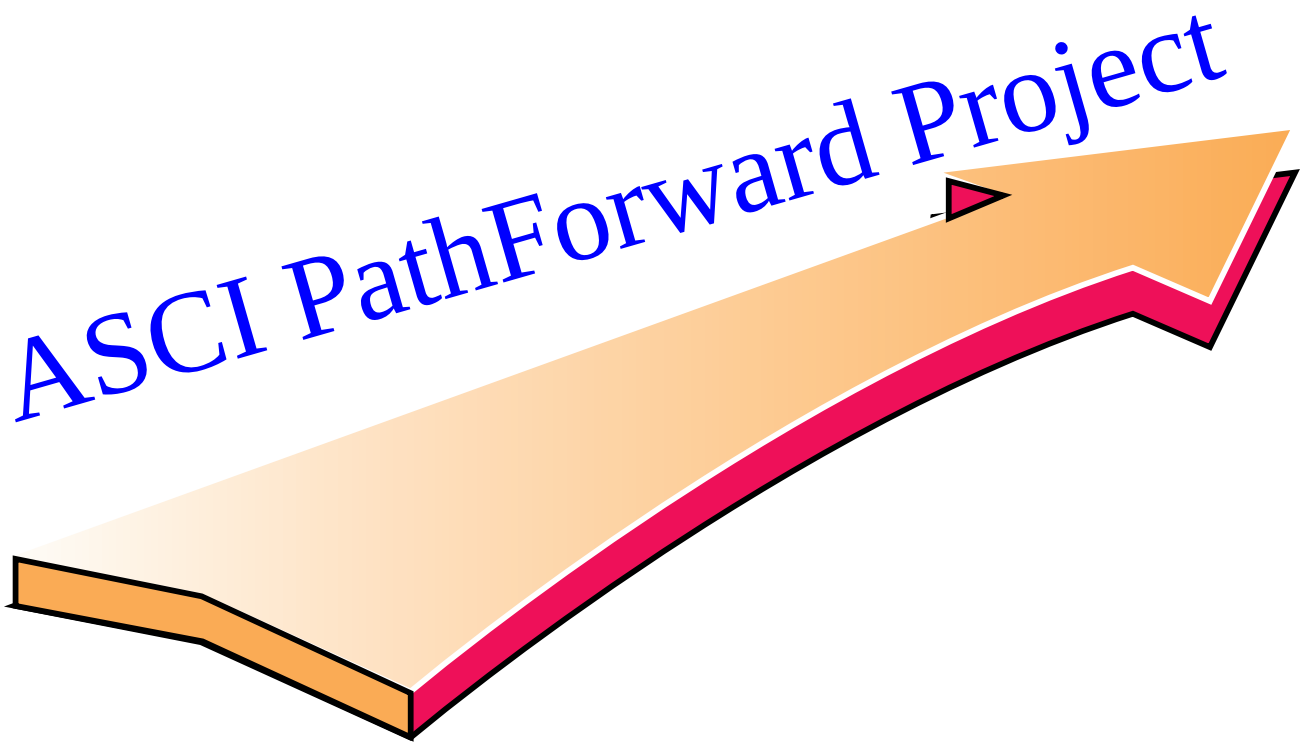
David Herr, LLNL

Gary Kent, DOE

Paul H. Smith, DOE

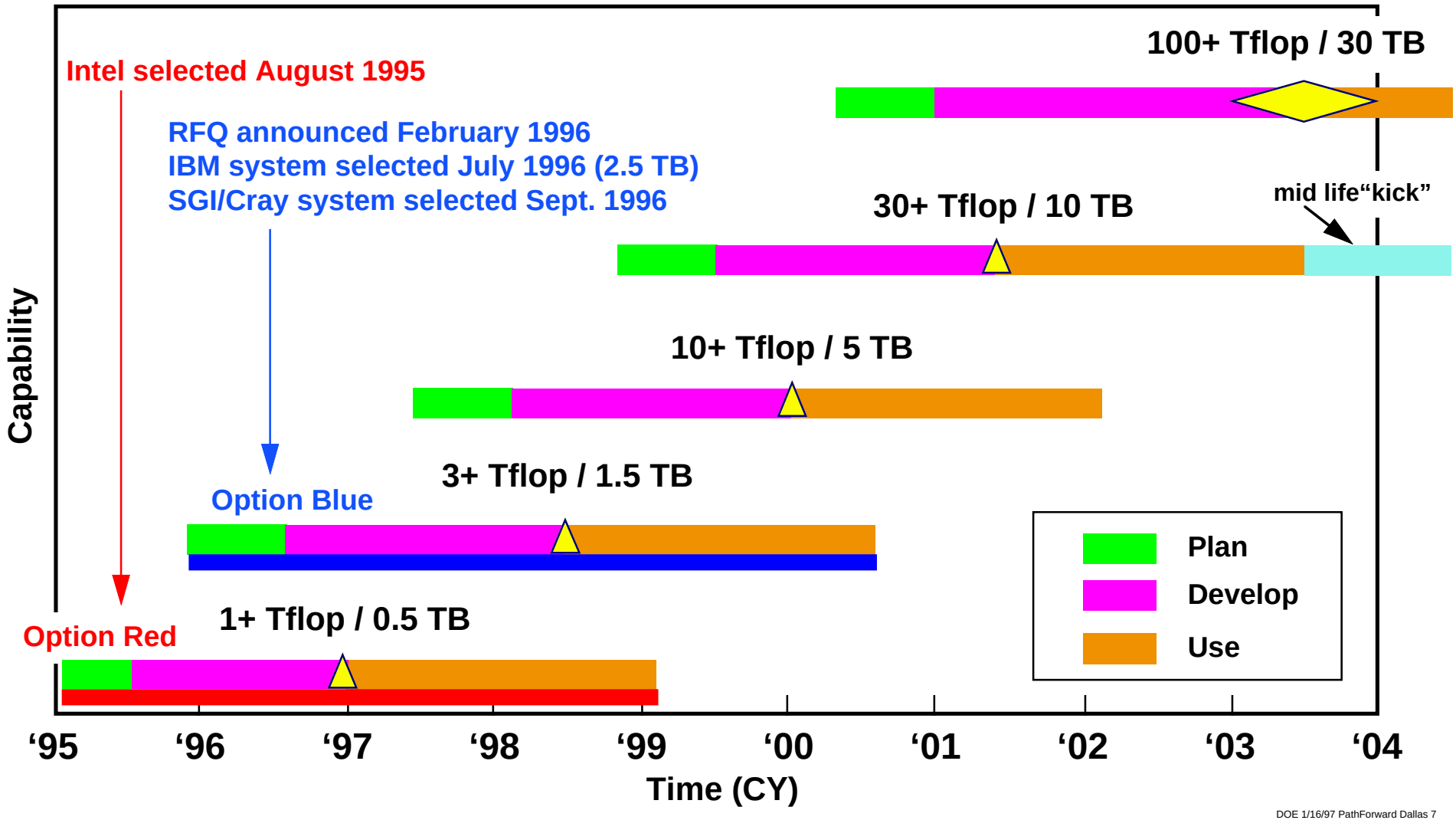
Dick Watson, LLNL

Karl-Heinz Winkler, LANL

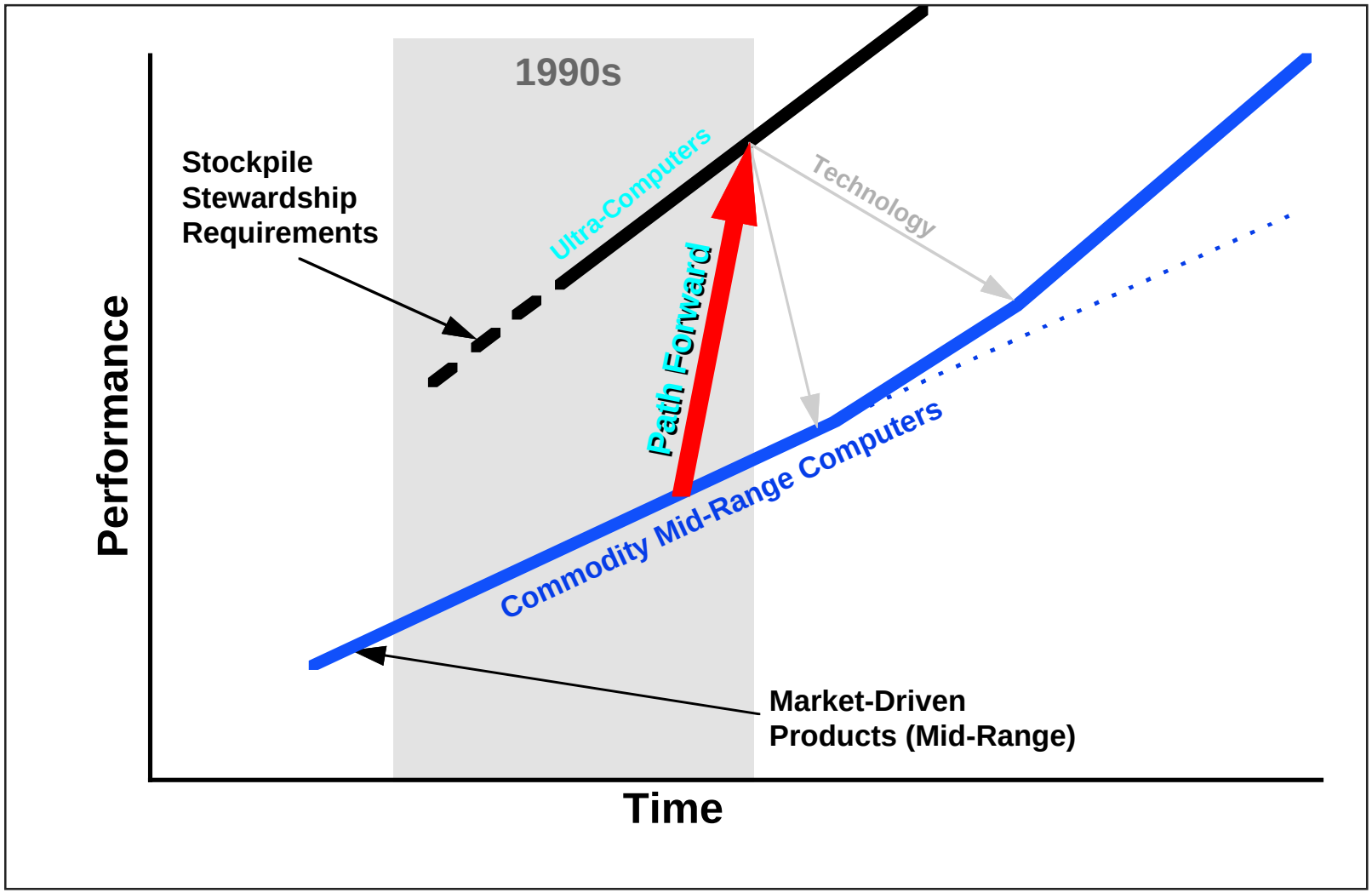

ASCI PathForward Project

ASCI Computing Systems Roadmap

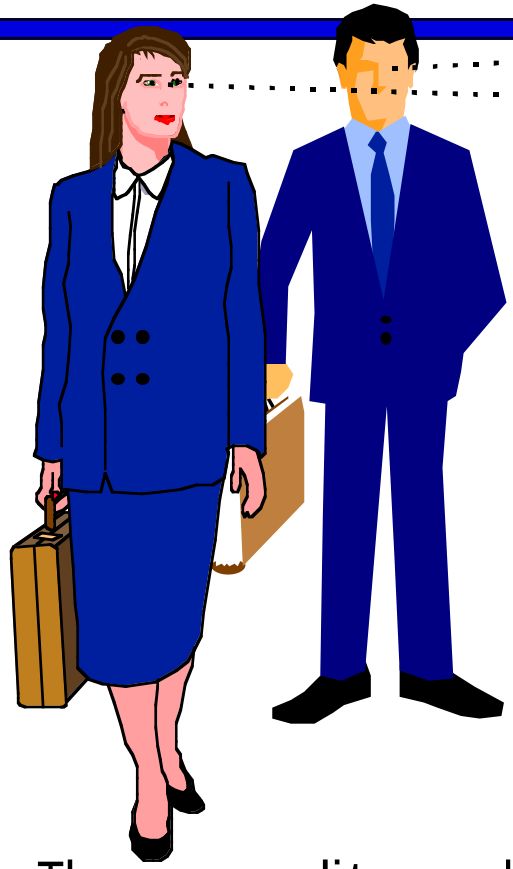
-- working with industry to reach unprecedented computer performance --



HEC Industry is not on the Path to Meeting Stockpile Stewardship Requirements



There is a new and fundamentally different source of High End Computing from US vendors--Commodity High End Commuting



commodity computing marketplace

- This is a business operated by people whose passion is making money and survival in a very competitive environment
- Performance achieved via scalability of commodity parts, particularly processors and memory
- The commodity computing marketplace is world-wide and very competitive

- The commodity marketplace is huge, 1000x or more, compared to government expenditures -- the government no longer dictates features
- Standard, commodity technology based on broader commercial requirements

Pathforward Strategy

Build future High End Computing (HEC) systems from building blocks (hardware and software) created to address the commodity computing marketplace.

concentrate investments (which are small compared to the commodity marketplace R&D) very carefully in those areas (hardware and software) which will enable commodity systems to be scaled-up to satisfy HEC requirements

PathForward Project: Development and Engineering Opportunities

- To achieve balanced, high-end, scientific computing at 10 TFLOPS and beyond,
- that can be made available on a time scale sufficiently rapid to impact future ASCI platform and subsystem procurements--notably 30 TFLOPS in the 2001 timeframe,
- that will accelerate extended capabilities of commercial product lines to market, and
- will not be created by commercial market forces in the time scale of interest.

PathForward Policy Objectives

- Establish industrial alliances with U.S. computer companies:
  - companies with a long-term stake in high-end scientific computing
  - PathForward links to company business plans rather than vice a versa
  - not overlapping university or national laboratory activities
- DOE investment in development and engineering projects:
  - ensure availability of essential integrating and scaling technologies required to create a well-balanced, reliable and production capable computing environment providing large scale, scientific compute capability from commodity building blocks, at processing levels of 10-to-30 TFLOPS in the late 1999 to 2001 time frame and 100 TFLOPS in the 2004 time frame
- Win-Win:
  - small perturbations in business plan
  - allows “ultra-scale” computing systems earlier

Achieving 100 TFLOPS ASCI Computational Performance

- Multiple high-performance commodity priced compute nodes, which represent regular commercial product lines and not special purpose designs;
- Hierarchical memory systems, including cache-only memory architectures and distributed shared memory systems with low-latency high-performance memory access;
- Commodity components and peripherals, such as DRAMs, disk and tape drives and controllers; supported by specially developed, essential integrating and scaling technologies, such as large-scale high-speed low-latency
- Communication fabrics, very high performance storage and parallel I/O systems, scaleable programming environments and operating systems, a universal programming paradigm, etc.

PathForward Process

- Request for Expressions of Interest 12/27/96
- Dallas Information Meeting 1/16/97
- Industry Expressions of Interest Due 2/18/97
- Dialogue with Industry Ongoing
- Bottom Line: Initiate Efforts ASAP

ASCI hardware requirements

Level	Effective Latency (CPU cycles)	Bandwidth (Random read/write)	Size
On-chip cache**, L1	2-3 ●	16-32 B/cycle ●	10^{-4} B/flop * ●
Off-chip cache**, L2 (SRAM)	5-6 ●	16 B/cycle ●	10^{-2} B/flop * ●
Local main memory (DRAM)	30-80 (15-30)	2-8 B/flop pk (2-8 B/flop sustained)	1 B/flop ●
"nearby nodes"	300-500 (30-50)	1-8 B/flop (8 B/flop)	1 B/flop ●
"far away nodes"	1000 (100-200)	1 B/flop (1 B/flop)	1 B/flop ●
I/O (memory disk)	10 ms ●	0.01-0.1 B/flop ●	10-100 B/flop ●
Archive (disk-tape)	Seconds ●	10^{-4} B/flop (0.001-0.01 B/flop)	10^{2} B/flop 10^{4} B/flop
User access	1/10 s (1/60 s)	OC3/desktop (OC12-48 /desktop)	100 users ●
Multiple sites	1/10 s ●	●	●

Compute engine

Interconnect

* Equivalent integer and floating-point data calculation rates are required.

** Cacheless systems with equivalent performance are fully acceptable.

Primary investment priority

Secondary investment priority

1996-1998 Situation

(1998-2000 Requirements)

Industry Trend

- Industry gets better at meeting requirements
- Industry gets worse at meeting requirements
- ● Industry continues to meet requirements

Current Hardware Trends (Commodity)

- industry, alone, will not meet ASCIrequirements -

Level*	Effective Latency (CPU cycles)	Bandwidth (Random read/write)	Size
On-chip cache**, L1	•	•	↑
Off-chip cache**, L2 (SRAM)	•	•	↑
Local main memory (DRAM)	↓	↓	↑
“nearby nodes”	↓	↓	•
“far away nodes”	↓	↓	•
I/O (memory disk)	•	•	•
Archive (disk-tape)	•	↓	↓
User access	•	↓	•
Multiple sites	•	•	•

Compute engine

Interconnect

Primary investment priority

Secondary investment priority

Industry Trend 1996-2000

↑ Industry gets better at meeting requirements

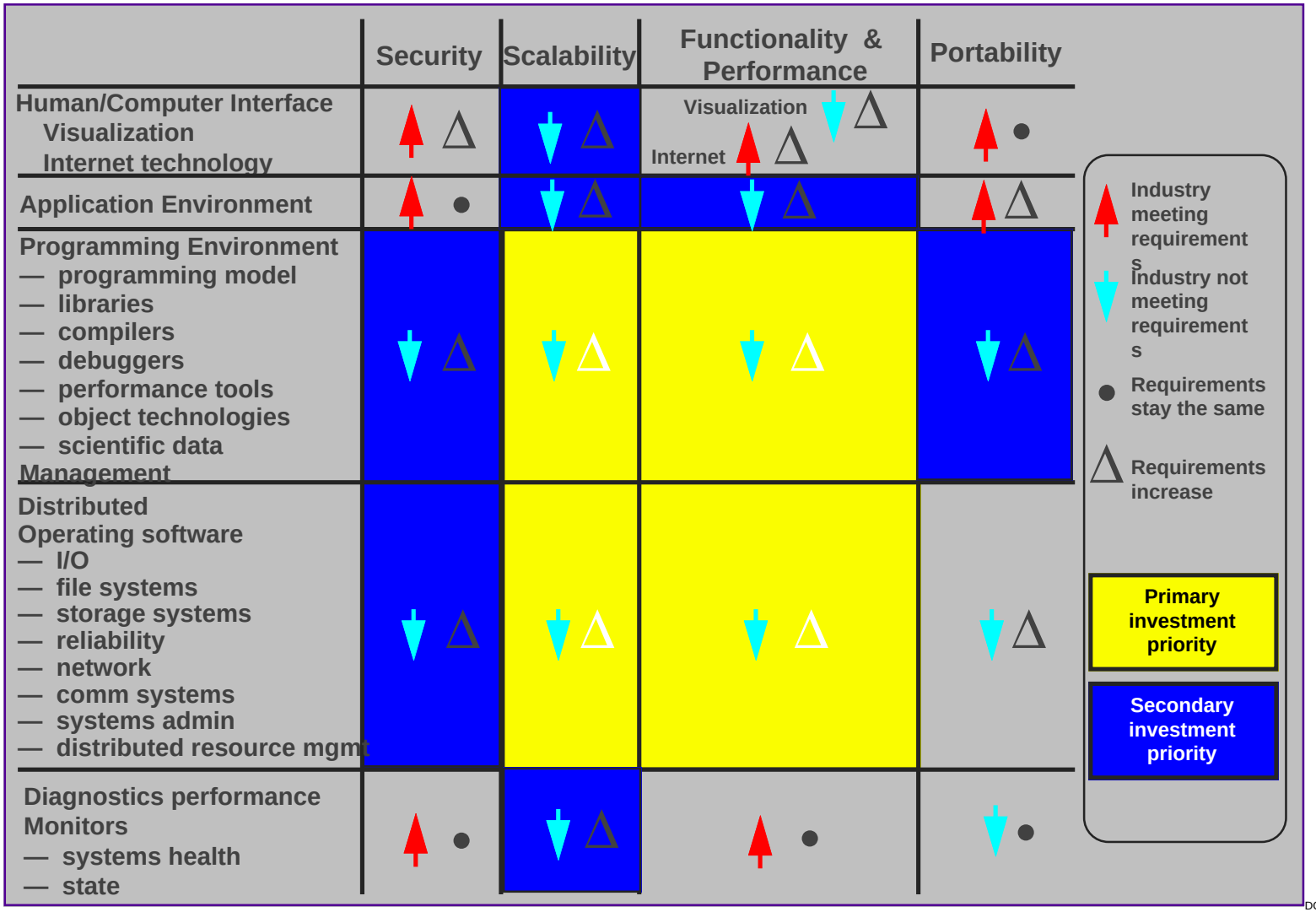
↓ Industry gets worse at meeting requirements

• Industry continues to meet requirements

* Equivalent integer and floating-point data calculation rates are required.

** Cacheless systems with equivalent performance are fully acceptable.

ASCI software requirements



Information Meeting

Information Meeting: January 16, 1996

Place of Meeting: Hyatt Regency
Dallas/Fort Worth Airport

Rooms Reserved: ASCI PathForward

PathForward HomePage

- ASCI PathForward Project Overview
- Request For Expression of Interest
- PathForward Project Description

  This document describes the goals and technical rationale of the PathForward Project.
- ASCI PathForward Project Meeting

  This document announces a Project Meeting being held on January 16, 1997.
- ASCI homepage

  This webpage contains links to other ASCI projects and issues.
- ASCI Program Plan

  This document provides a global overview of the ASCI program.
- Frequently Asked Questions (FAQ)

ASCI is Focused on the High End of Computing

- Accelerate High-Performance Computing through Multiple-Partnerships
- Develope Software to take full advantage of hardware capabilities
- Maintain affordability by acclerating existing industry technology trends
- Provide balanced 100 TFlops computing platforms by 2003/4

Balanced "Ultra-Scale" Computing

- Balanced systems at the 10 to 100 TFLOPS scale
  places stringent requirements on the internode interconnect, I/O, and storage subsystems.
- ASCI balanced systems ratios:
  1 TeraFLOPS peak performance /
  1 Terabyte memory size /
  50 Terabyte disk storage /
  16 Terabyte per second cache bandwidth /
  3 Terabyte per second memory bandwidth /
  0.1 Terabyte per second I/O bandwidth /
  10 Gigabyte per second disk bandwidth /
  1 Gigabyte per second archival storage bandwidth /

.In scaling up over several orders of magnitude, these ratios will be difficult to achieve.

PathForward HomePage

- ASCI PathForward Project Overview
- Request For Expression of Interest
- PathForward Project Description
  This document describes the goals and technical rationale of the PathForward Project.
- ASCI PathForward Project Meeting
  This document announces a Project Meeting being held on January 16, 1997.
- ASCI homepage
  This webpage contains links to other ASCI projects and issues.
- ASCI Program Plan
  This document provides a global overview of the ASCI program.
- Frequently Asked Questions (FAQ)